Unitization of Familiar Letter Patterns

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for Graduation with Distinction in Psychology in the Undergraduate Colleges of The Ohio State University

By

Sarah Poston

The Ohio State University

2006

*****

Project Advisor: Dr. Neal F. Johnson, Department of Psychology

Abstract

Research on letter-priority effects has demonstrated that for pronounceable nonwords, non-native English speakers are faster at identifying individual letters than they are at identifying entire non-words. Conversely, for pronounceable *words*, subjects identify the entire word faster than they can pick out any one individual letter. What has not yet been explored is the way in which we process unitized familiar letter patterns, or acronyms. This experiment compares the reaction time to identify a predesignated target, the first letter or the entire array, of familiar and unfamiliar letter arrays. Familiar arrays show word-priority effects similar to familiar words. Unfamiliar acronym arrays do not show letter priority effects for native English speakers.

Although the study of word processing began in the early 1900's with Edmund Huey's book on reading, we still know relatively little about the mechanisms that allow us to see a word, recognize and encode it, and realize that it has some significance to our task. Researchers have come up with two models with which to explain how we encode words. The two models debate whether we read by recognizing individual letters and then assembling them to form words, or whether we view the entire word as one unit without the extra step of letter identification.

The theory of holistic processing, which assumes that we skip letter identification for words, has had data to support it since the late 1800's. Coggeshall, in her 1999 thesis, cites Cattell's work (1886, in Huey 1908), which revealed that the time needed to recognize a word and a single letter (outside the context of a word), were virtually the same, and that word length had no effect on reaction time. Johnson (1975a) also demonstrated no word length effects in discussing his "pattern-unit" model of word identification. He compared words from 3 letters to 8 letters and found no significant difference in the amount of time it took subjects to respond to the items. This ease with which subjects are able to recognize words as one unit, has been called the word-priority effect (Sloboda, 1976; Johnson, 1975a). This effect is characterized by shorter reaction times to words, than to individual letters within words, when subjects are asked to compare two displays and make a matching decision. Healy and Drewnoski (1983, in Johnson 1991) also demonstrated that shorter words like *and* and *the* conceal their letters. Subjects, when presented with short, very frequent words did not consistently detect typographical errors. The word level code takes priority and interferes with the letter level code. Johnston and McClelland (1980 in Johnson 1991) used a hierarchical model to assert that memory is organized in a top-down manner – the reason that words are activated before individual letters. It is assumed that word targets are the default (i.e. word encoding is holistic). Only if the subjects are asked to identify letter information, do they take steps to overcome the default holistic memorial code.

Using the aforementioned studies as stepping stones, two models were developed to explain the process of word and letter identification. The first model asserts that words are viewed holistically. When we experience a stimulus such as an English word, we see an image of that word (e.g., table) in our sensory memory. Sensory memory sends that image to an encoder which accesses long-term memory and searches for a pre-existing memory code. We have previously seen and memorized the word "table" and are thus able to access it with great speed. From there, the memory code is assigned to the display and it travels to a tester that decides if it matches what we have seen. If it passes the test, the memory code continues on to our working memory where we may work with the word in our memory systems (Johnson, 1991). (See Conceptual Model 1)

The second model maintains that nonwords are processed on a letter-by-letter basis. For example, if the image is of the nonword "vable," we will not have a pre-existing memory code to access in long-term memory and the image will be tossed back and forth between the encoder and tester until the two finally concede that there will never be a match. The nonword is then *parsed* into letter units and each letter is encoded individually. We have pre-existing memory codes for all the letters of the alphabet, and we are therefore easily able to move all the letters into working memory where we may reassemble the nonword (Johnson, 1991). This is, in essence, the letter-priority effect. (See Conceptual Model 2)

Coggeshall (1999), building on the theory that word-level information is available before individual letter information for words, investigated word- and letter-priority effects within the realm of pronounceable nonwords (e.g. glock). Nonwords do not have the advantage of a pre-existing memory code for identification and are thus not processed as easily. Nonwords, without a pre-existing memory code, require extra processing to parse the word into its individual letter components, identify each letter and its phonological significance to the nonword, and finally reassemble it to create a recognizable, pronounceable nonword.

Using the holistic model as a basis for understanding word processing, Coggeshall (1999) tested the hypothesis that if subjects were forced to parse a display during processing by being asked about letter information, they could more quickly respond to a yes/no question regarding individual letter information in a nonword than they could about the entire nonword as a whole. When Coggeshall analyzed her data she found that this was not the case. English speakers are very skilled in phonological computation since they are taught from childhood to sound out words that are new or difficult to pronounce. Instead of finding a letter-priority effect for the pronounceable nonwords, there was virtually no difference between the time it took to identify a single letter within a nonword or to identify the entire nonword itself. She explained the lack of a disparity between letter and entire array searches by hypothesizing that native English speakers were so skilled at the assembly of pronounceable nonwords that the test was not sensitive enough to detect any slight difference that might exist.

As a result of that idea, Coggeshall (1999) became interested in the phonological computation of subjects whose primary language is not English. Subsequent research focused on individuals whose primary language was one without an alphabetic language (e.g.. Chinese, Japanese and Korean students). These subjects would not be as familiar with grapheme-to-phoneme conversion rules (computing the image of a letter into its phonological sound), in English, and thus processing alphabetically would not be as automatic. Contrary to expectations with regard to words, these Asian subjects had an identical word-priority effect to the English-speaking students. However, when the Chinese, Japanese and Korean students were exposed to pronounceable nonwords, they displayed a huge letter-priority effect. They were approximately 50 milliseconds faster at identifying an individual letter in a nonword than the entire nonword itself. Coggeshall's results confirmed that pronounceable nonwords do, in fact, show a letter-priority effect.

In summary, non-Asians displayed the word-priority effect in that they made matching decisions more quickly about words than about individual letters within words. They exhibited no letter-priority effect; the time to identify nonwords and letters within nonwords was almost indistinguishable. Asian subjects displayed an identical word-priority effect to their English-speaking constituents, but did exhibit a letter-priority effect. They were faster at identifying individual letters within a nonword than identifying the entire nonword itself.

One type of nonword letter array that differs from a typical nonword is an acronym. It also differs from words and nonwords in its pronounceability. Coggeshall (1999) based her theory on the assumption that the word-priority effect is dependent upon lexical access and an orthographic representation, meaning that there is a pre-existing memory code for the word or individual letters, and that it follows the rules of the English language and is pronounceable. Acronyms are short letter patterns that stand for a longer phrase, and many of these have become integral parts of the English language. Most individuals, when referring to the Federal Bureau of Investigations use the acronym FBI. It would be fairly unlikely to encounter an individual in the United States who does not understand the significance of those three letters. In other words, most US residents would have a pre-existing memory code for the letter array "FBI." Contrary to the above studies, FBI is neither a word, nor is it pronounceable and thus it poses the question as to whether a lack of orthographic regularity is significant enough to inhibit a letter pattern-priority effect or acronym-priority effect. That is, will the fact that acronyms are unpronounceable cause letters within an array to be identified more quickly than entire array. If orthographic regularity is not required for word processing, subjects should be faster at array identification.

Unfamiliar arrays of letters may or may not show a letter-priority effect (Johnson, 1991; Coggeshall, 1999). This builds on previous research that demonstrates that pronounceable nonwords with which we are not familiar, show a strong letter-priority effect for non-native English speakers. Pronounceable nonwords only exhibit the expected letter-priority effects for non-native English speakers since English speakers are so skilled at grapheme-to-phoneme conversion that there is no difference between letter detection and letter pattern detection. This same unitizing skill may be used to quickly compute codes for unfamiliar acronyms and could result in no letter-priority effect.

Conversely, familiar letter arrays should exhibit a letter-pattern priority effect or acronym-priority effect. Presumably, letter arrays with which we are familiar (e.g. FBI) will be moved into working memory as a unit. In essence, this means that the difference between array search and letter search for very familiar acronyms will be larger than the difference between array and letter search for unfamiliar acronyms. In addition, more time will be required to answer NO than YES in this decision making task because participants are generally expected to perform a double check when there is no match between the stimulus and response.

Method

In order to explore this issue, three types of displays were used. Reaction times (RT) to 36 three-letter, very familiar, non-pronounceable acronyms were compared with the RT to 36 three-letter, semi-familiar and 36 three-letter, unfamiliar arrays, and all three types appeared equally as often. The experiment used a two (task type: letter as target or array as target) by three (display type: very familiar, semi-familiar and unfamiliar acronyms) by two (response type: yes vs. no) within subject design.

Participants

Sixty-four participants were recruited from Psychology 100 classes at The Ohio State University. Linguistic background was not a factor used in selecting subjects, although it may be assumed that since the students were attending an English speaking university, they must have some familiarity with the English language. If non-native English speakers were given this same task, it would be assumed that they would show a stronger letter-priority effect for unfamiliar acronyms than their native English-speaking counterparts.

Procedure

Participants were given verbal instructions about the task and the type of display, and the same information was displayed on the computer screen at the beginning of each trial. Prior to the display, subjects were told they would be answering yes/no questions regarding the letter or array information to be displayed on the screen. These arrays appeared in the middle of the screen on a Televideo 920C Terminal. Participants pressed the z-Key, marked Y, or the ?-key, marked N, to respond to the questions. Before each letter array appeared on the screen, subjects were asked to match the array with a predesignated target. The targets were either the first letter or the entire array of letters. The pressed "enter" one time to display the target on screen. While the target was displayed, the experimenter named the target out loud. Then, the experimenter pressed "enter" to advance the display to an X in the middle of the screen followed by a dot for 500 milliseconds (msec) and finally the letter array appeared in place of the dot. The display was terminated when the participant made the YES or NO decision and pressed the Y or N key. Response latency and accuracy were recorded for data analysis.

Results

The median reaction time was calculated for each participant in each condition. The latency data are displayed in Table 1. There was a significant main effect of display type, $F(2, 126) = 14.83$, $p < .001$. The means for very familiar acronyms, semi-familiar acronyms and unfamiliar acronyms were 590 msec, 595 msec, and 599 msec respectively. There was also a significant main effect of task type, $F(1, 63) = 6.26$, $p <$

.015. The mean for the array search task across all levels was 574 msec, and for the letter search task across all levels the mean was 615 msec. Response type also yielded a significant main effect, $F(1, 63) = 130.3$, $p < .001$, with YES responses faster than NO responses. YES responses averaged 582 msec and NO responses averaged 618 msec; a phenomenon that is commonly found in reaction time studies, and it seems to be the result of a double-check that occurs when targets and displays do not match.

The critical interaction for this experiment, between display and task type was significant, $F(2, 126) = 5.14$, $p < .009$. This interaction is displayed in Figure 1. The acronym displays yielded a significant acronym-priority effect; decisions about the array were made more quickly than decisions about individual letter information across all display types, very familiar, semi-familiar and unfamiliar acronyms. However, for letter searches, unfamiliar acronyms did not have the quickest reaction time, shedding doubt on the presence of a letter-priority effect in unfamiliar acronyms. An analysis of the very familiar and semi-familiar items indicated that the acronym-priority effect was significant, $F(1, 63) = 19.09$, $p < .001$, but the difference between the very familiar and semi-familiar displays was not significant, $F(1, 63) = 2.57$, $p > .05$. In addition, the display-by-task interaction was not significant, $F > 1.00$ indicating that the acronym-priority effect was approximately equal for those two conditions. There was a letter-pattern priority effect for the unfamiliar items as well, $F(1, 63) = 8.04$, but it was clearly smaller than for the other two display types (35 msec vs. 42 and 44 msec for very familiar and semi-familiar acronyms), and that resulted in the significant display-type by task-type interaction in the overall analysis.

Other significant two-way interactions were found between display type and response type, $F(2, 126) = 5.62$, $p < .006$, and between task type and response type, $F(1, 63) = 5.18$, $p < .03$, and they are displayed in Figures 2 and 3 respectively. From the very familiar acronyms to the unfamiliar acronyms, there is a larger increase in latency for the YES responses than there is for the NO responses. Figure 3 displays that there is a larger difference between letter and array reaction times for No responses than for YES responses.

Finally, the three-way interaction between display type, task type, and response type was also significant, $F(2, 126) = 4.03$, $p < .02$. For YES responses, the latency differences between array and letter search were larger in the very familiar and semi-familiar acronym conditions than the difference between array and letter search in the unfamiliar acronym condition. For NO responses, the differences were virtually the same. This interaction is displayed in Figures 4 and 5.

The error data are displayed in Table 2. Only the main effect of response type was significant, $F(1, 63) = 4.79$, $p < .04$. The main effects of display type and task type were not significant, $F < 1.00$ for both types. Neither the two-way interactions between display type and task type, nor between display type and response type were significant. Both had an $F < 1.00$. In addition, the task by response interaction was not significant, $F(1, 63) = 1.20$. Finally, the three-way interaction between display type, task type and response type was not significant, $F(2, 126) = 1.40$.

The correlation between reaction time and errors was -.35, and while this could present evidence for a speed-accuracy trade-off, a correlation of .55 is required for significance when n=12. It may also be assumed that the -.35 occurred because of the interaction between YES and NO answers. When the correlation was computed using only data from the YES responses, which is really the important factor here, it becomes +.11. This statistic is also not significant, but it also is not negative, and we can conclude that there was little or no speed-accuracy trade-off.

Discussion

As expected, an acronym-priority effect was found for very familiar and semi-familiar acronym displays. It took longer for subjects to respond when the target was an individual letter within the acronym than when the target was the entire array. It also took longer to identify individual letters in the unfamiliar displays, and there was no letter priority effect. (See Figure 1). When response type and task type are compared, it reveals that array searches were performed more quickly than letter searches for all displays types (See Figure 3).

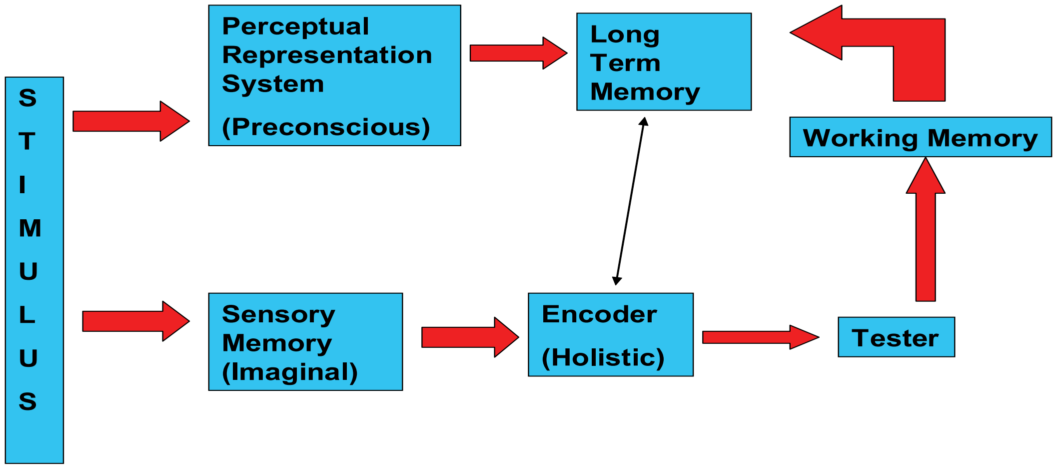
The display by response interaction confirmed that YES decisions are made more quickly than NO decisions, because subjects will terminate their search when they find the target. If the target is not found, subjects will double-check before making a decision (See Figure 2).

Examination of the three-way interaction between task, display and response types confirms that letter searches took significantly longer than array searches in the very familiar and semi-familiar display types. For the unfamiliar acronym displays, the difference between letter and array search was not as large (See Figures 4 and 5).

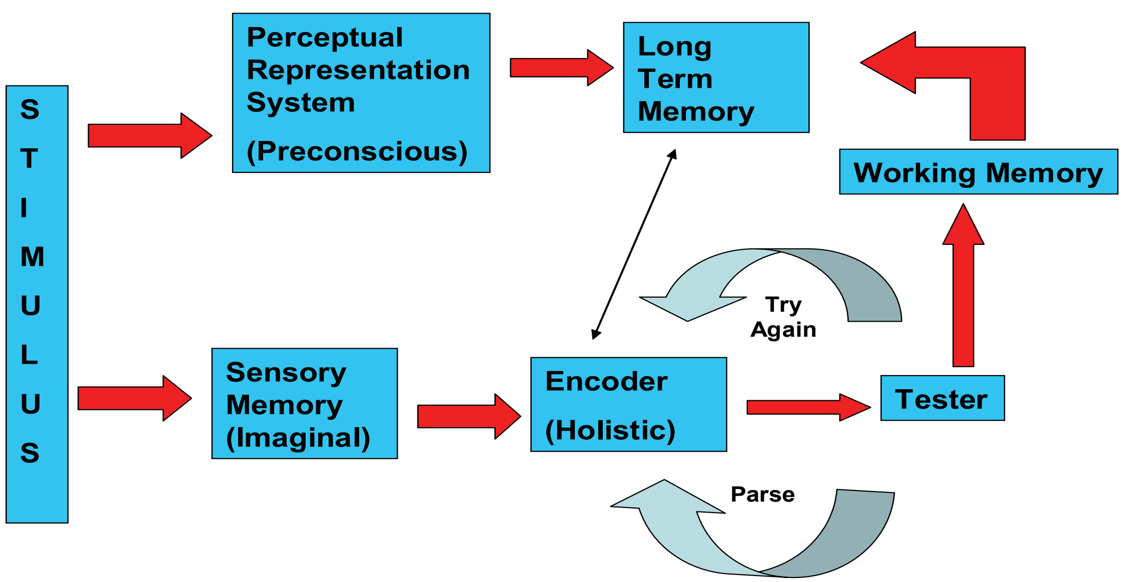
One possibility is that this phenomenon occurred due to the theoretical perspective that English speakers are very skilled at grapheme-to-phoneme conversions. In the very familiar and semi-familiar arrays, letter searches take more time than array searches due to the extra step of parsing the information from the holistic code into individual letter units. As was expected with English-speaking subjects, in the unfamiliar display condition, letters were not identified more quickly than entire arrays. In Coggeshall's (1999) experiment, letters within nonwords and nonwords were identified in virtually the same amount of time. In this experiment, there was a difference in the amount of time it took to identify a letter within an unfamiliar acronym and the entire acronym. The time to identify the entire array was 581 msec and to identify a letter within the array it took 616 msec.

Perhaps instead of utilizing the letter information that was immediately available to them, subjects attempted to unitize the unfamiliar patterns since they had no pre-existing memory code. This step takes longer than for very familiar and semi-familiar acronyms, since they already have a pre-existing memory code for those items. After the unfamiliar acronym is unitized, it can be quickly parsed into individual letter units since the unitized code has only been reinforced during one trial. The combination of two processes, unitization and parsing for unfamiliar letter patterns, takes slightly longer than for one process, parsing, in the very familiar and semi-familiar acronym displays, since these displays have such strong pre-existing memory codes.

This study is important because it contributes to the need for an investigation into the computational processes involved in recognizing words to enhance the way in which we teach our children to read. For example, sufferers of dyslexia may not able to unitize patterns as easily as others, and are slowed down by individual letters within the word, rather than just skipping over them as most readers do. A subsequent comparative analysis between skilled and dyslexic readers performing the same letter and word search tasks would provide insight into the depth of deficiency and some of the mechanisms at work.



Conceptual Model 1. Processes employed in memory searches for words or very familiar acronyms (based on Johnson, 1991).



Conceptual Model 2. Processes employed in memory searches for nonword or unfamiliar acronyms (based on Johnson, 1991).

ARRAY-SEARCH TASK

Display Type	Yes	No	Mean
Acronyms			
Very Familiar	553	585	569
Semi-Familiar	558	588	573
Unfamiliar	571	591	581
Mean	561	588	574

(Response Type: Yes / No / Mean)

LETTER-SEARCH TASK

Display Type	Yes	No	Mean
Acronyms			
Very Familiar	580	642	611
Semi-Familiar	587	647	617
Unfamiliar	581	651	616
Mean	583	647	615

ACRONYM/LETTER SEARCH DIFFERENCES

Display Type	Yes	No	Mean
Acronyms			
Very Familiar	27	57	42
Semi-familiar	29	59	44
Unfamiliar	10	60	35
Mean	22	59	41

Table 1: Latency data (in msec).

ARRAY-SEARCH TASK

	Response Type		
	Yes	No	Mean
Display Type			
Acronyms			
Very Familiar	5.4	5.2	5.3
Semi-Familiar	5.2	5.2	5.2
Unfamiliar	5.2	5.2	5.2
Mean	5.26	5.2	5.23

LETTER-SEARCH TASK

	Response Type		
	Yes	No	Mean
Display Type			
Acronyms			
Very Familiar	5.5	5.3	5.4
Semi-Familiar	5.3	5.2	5.25
Unfamiliar	5.3	5.1	5.2
Mean	5.37	5.2	5.29

Table 2: Error data (in percentages).

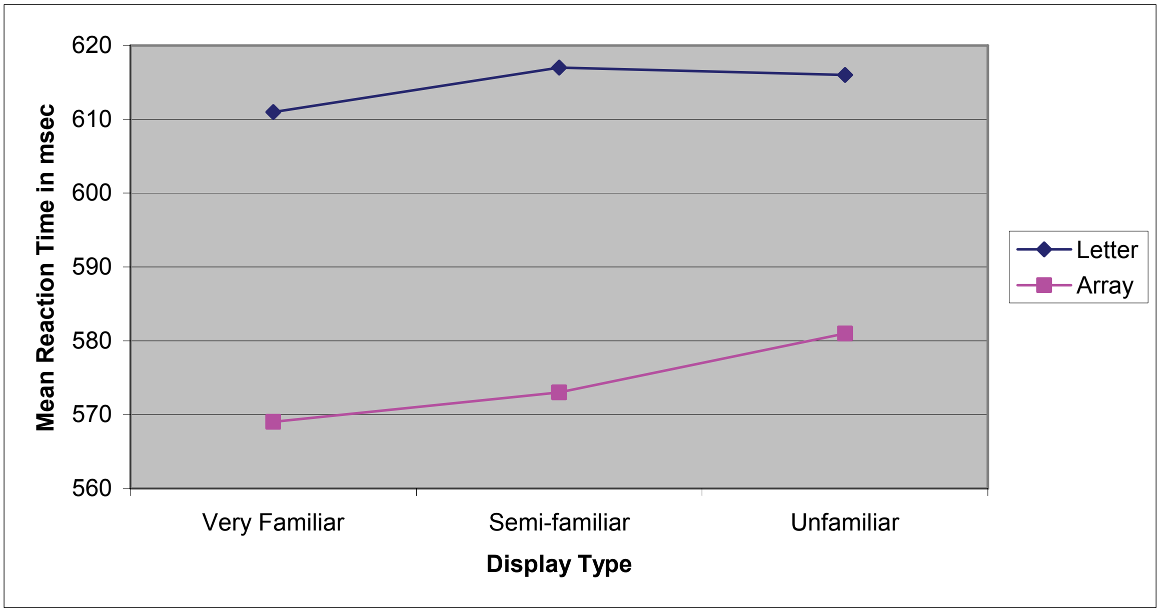


Figure 1: Task-by-display type interaction

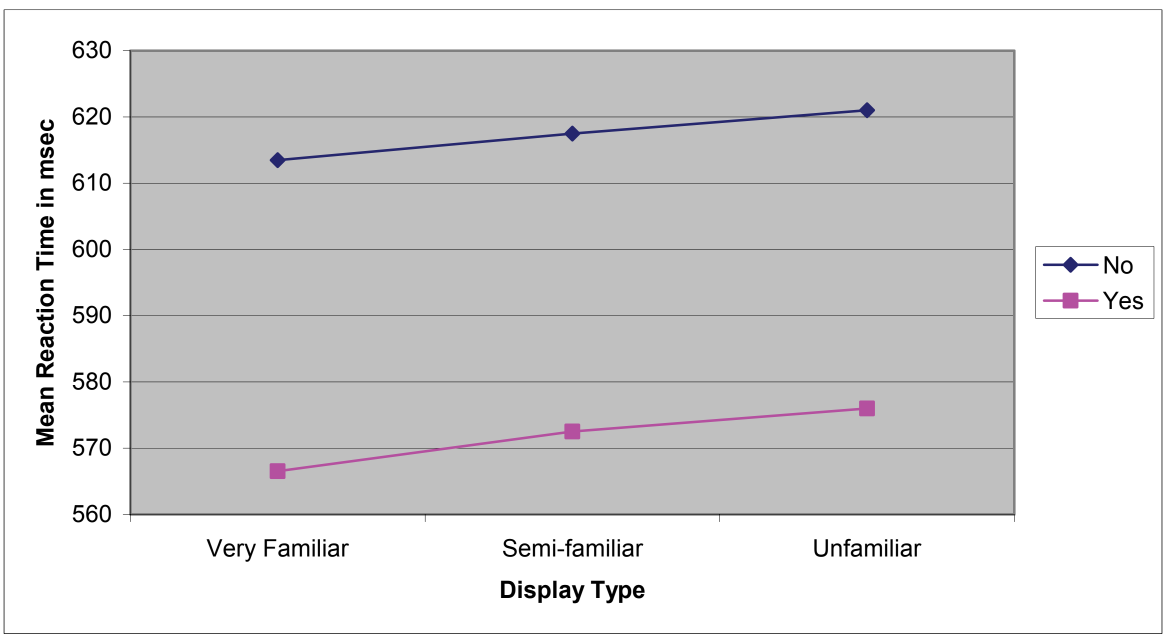


Figure 2: Display-by-response type interaction

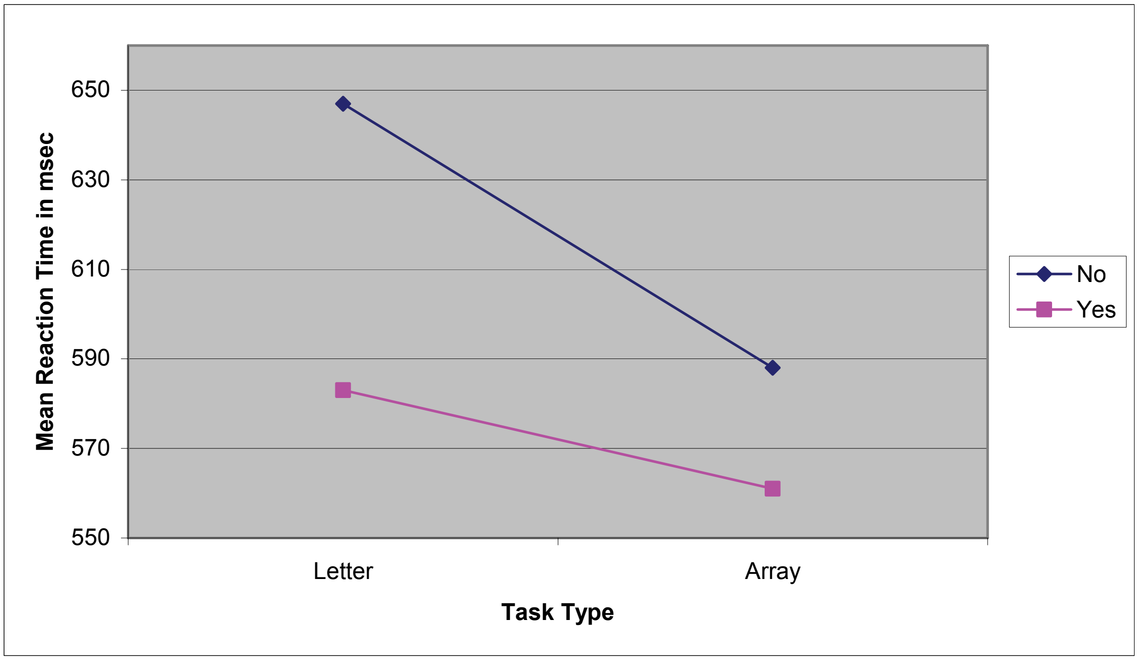


Figure 3: Response-type-by-task type interaction

YES RESPONSES

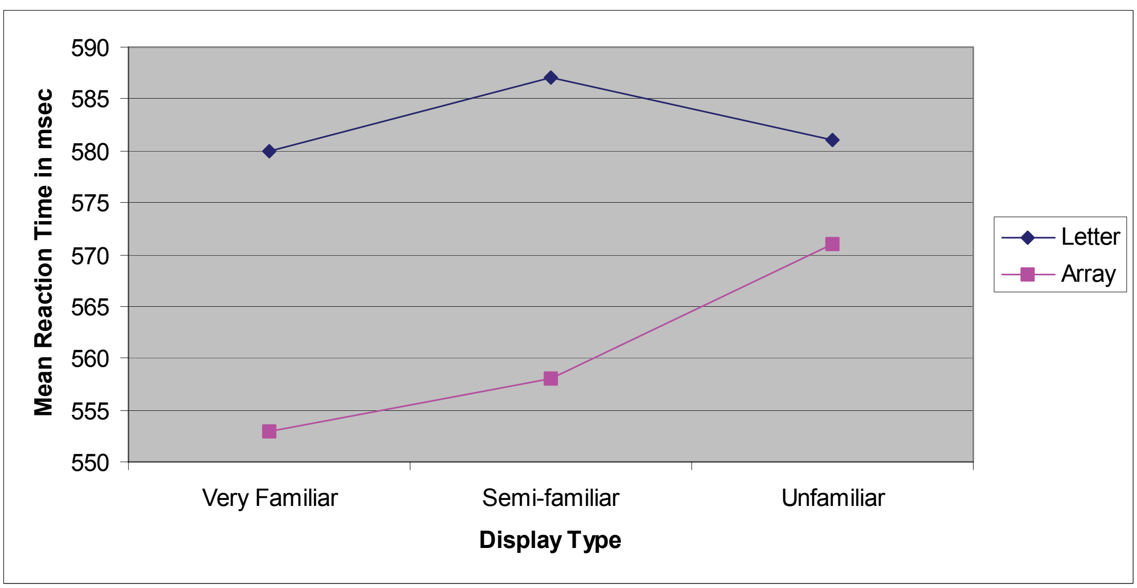


Figure 4: Task-by-display-by-response type interaction

NO RESPONSES

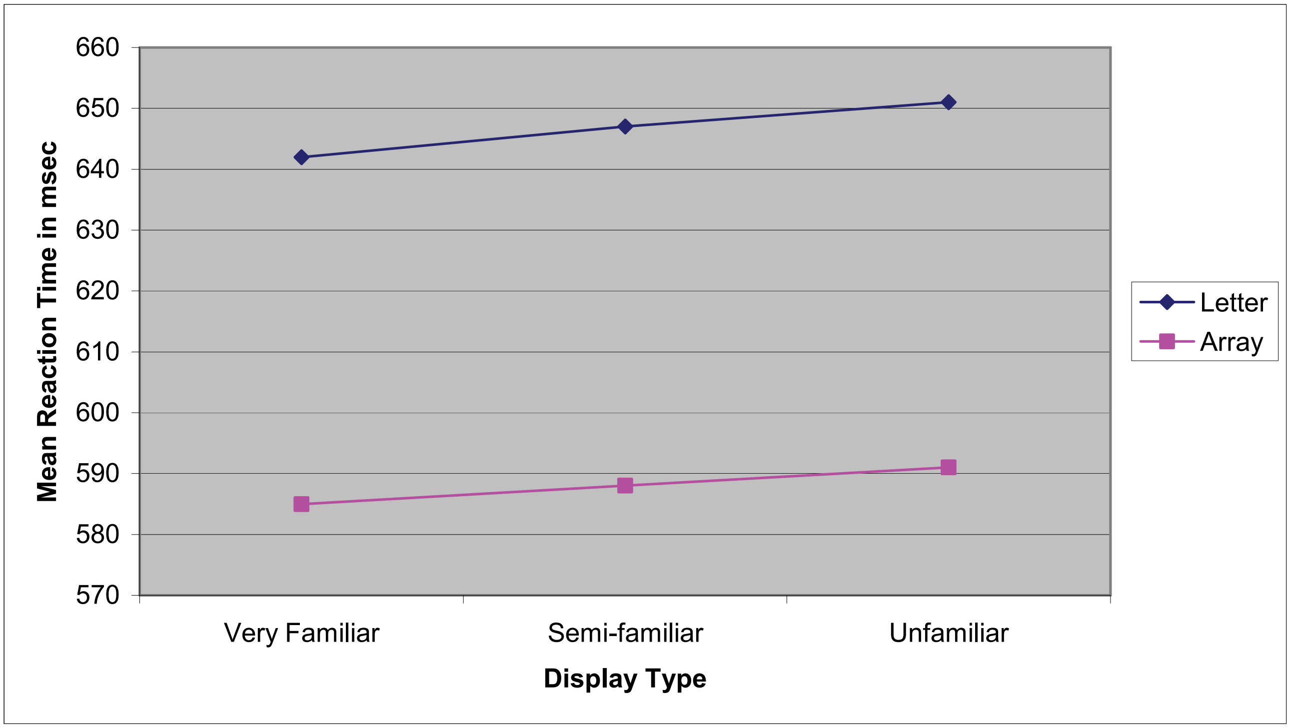


Figure 5: Task-by-display-by-response type interaction

APPENDIX A

Stimulus lists

List A, Very Familiar Acronyms. Each row lists target/display pairs.

1. ABC, PIV	19. NBA, ZIC
2.AOL, DER	20. NBC, WUD
3. ATM, NEF	21. NFL, CAG
4. BLT, JOR	22. NHL, BOJ
5. BMW, KIR	23. OSU, FID
6. CBS, LUN	24. PBS, LIR
7. CIA, TUL	25. PDA, BIS
8. CNN, GAW	26. PhD MAZ
9. DNA, HEG	27. STD, VAB
10. DUI, POQ	28. TBS, DOK
11. FBI, JAL	29. TLC, HIB
12. FDA, CIZ	30. IRS, NUF
13. HBO, FES	31. UPS, NAX
14. HIV, MUJ	32. USA, QOB
15. KFC, QUN	33. VCR, GEZ
16. PGA, WIB	34. VHS, KAL
17. MLB, SIZ	35. VIP, PEB
18. MTV, RIH	36. ESP, TIG

List B, Semi-familiar Acronyms. Each row lists target/display pairs.

1. BMV, DAF	9. EST, VID
2. CEO, RUD	10. ETA, FID
3. CPA, GEG	11. FCC, KUV
4. CVS, MOG	12. GED, PIM
5. DOA, CUX	13. GPS, LAV
6. DSW, HAB	14. IBM, ZER
7. DWI, BEK	15. IRA, JOM
8. EKG, DOJ	16. LCD, MOT

17. LSD, PAG
18. MCI, QUK
19. MIA, TOF
20. MLS, WIK
21. MRI, ZAD
22. PGA, YID
23. POW, RUR
24. QVC, WOR
25. RCA, YUT
26. TWA, HUC
27. UDF, JIR
28. UPN, BIK
29. URL, MOT
30. SBC, NUJ
31. VFW, CUG
32. LTD, SAR
33. DEA, RIN
34. AFL, ROV
35. AEP, NOC

List C, Unfamiliar Acronyms. Each row lists target/display pairs.

1. GFE, SUL
2. OLJ, SAZ
3. XCA, JOF
4. NRU, GAT
5. ITG, HUD
6. HWA, BUJ
7. LQE, DIR
8. BVO, CER
9. SZA, KIG
10. IJM, FOZ
11. SPV, QUM
12. NMO, TER
13. TFI, DUS
14. QPE, BOV
15. BHA, LUT
16. OGX, VEZ
17. LXA, ZOK
18. BJI, MEZ
19. ASZ, NUD
20. PYO, WEK
21. VGU, LAR
22. CJA, YIG
23. MXE, CUJ
24. UWB, RIC
25. OJN, GUL
26. RNE, PAG
27. IYB, RUV
28. PFA, MIG
29. NGU, SOM
30. LHE, FAW
31. QRI, JAT
32. DGA, PIW
33. HVA, NUR
34. BXI, HOD
35. RQU, TAM

References

Coggeshall, L. (1999) *On the role of letters in the processing of words and nonwords*. Unpublished master's thesis, The Ohio State University, Columbus, OH.

Huey, E. B. (1908). The psychology and pedagogy of reading. Cambridge, MA: MIT Press.

Johnson, N. F. (1974) A pattern-unit model of word identification. In D. Laberge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp.91-125). Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson, N.F. (1991) Holistic models of word recognition. In R. R. Hoffman & D. S. Palermo (Eds.), *Cognition and the symbolic processes: Applied and ecological perspectives* (pp. 79-93). Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson, N.F., Childers, T.L., Jankowski, S., Miller, B., Gonsman, V.L. & Seifert, L.S. (2004) *A cohort model of lexical processing: A conceptual framework for neighborhood effects and supporting data*. Unpublished manuscript, The Ohio State University, Columbus, OH.

Sloboda, J. A. (1976). The locus of the word-priority effect in a target-detection task. Memory and Cognition, 5 (3), 371-376.